A NOVEL APPROACH FOR MULTI-FACILITY CAPACITY EXPANSION AND CONTRACTION PROBLEMS UNDER UNCERTAINTY

A Thesis

by

RICHARD C. ALLEN

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Douglas Allaire
Committee Members,	Mahmoud El-Halwagi
	Shima Hajimerza
Head of Department,	Andreas A. Polycarpou

August 2018

Major Subject: Mechanical Engineering

ABSTRACT

In this work we present a novel solution strategy for the development of processing facilities for shale natural gas fields that is more sustainable and cost effective. The foundation of our methodology is to utilize flexible facilities that are comprised of transportable modular processing units as opposed to fixed capacity permanent facilities. These processing units can be purchased for and reallocated between facilities at points within a discretized finite planning horizon. This in turn allows the capacity of each facility the ability to adapt through either expansion or contraction with respect to the uncertain and dynamic influent flow-rate. We illustrate that our methodology as applied to a random set of test cases with three or more facilities is at worst identical to utilizing permanent facilities and on average at least 12% more cost effective.

DEDICATION

To my mother, my father, and my late grandparents.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a thesis committee consisting of Dr. Douglas Allaire of the Department of Mechanical Engineering [Chair] and Dr. Shima Hajimerza of the Department of Mechanical Engineering and Professor Mahmoud El-Halwagi of the Department of Chemical Engineering.

All work conducted for the thesis was completed by Richard C. Allen independently.

Funding Sources

Graduate study was supported by a teaching assistantship provided by Texas A&M University.

NOMENCLATURE

$\mathcal{L}$	Set of facilities where a skid can be located
$\mathcal{B}$	Set of preexisting processing skids
$\mathcal{C}$	Set of skid sizes/technologies available from suppliers
$\mathcal{N}$	Set of time periods
$\mathcal{F}$	Is a collection of sets, where $f \in \mathcal{F}$, represents the set of nodes that represent a forecast
$\mathcal{N}$	The nodes in the skid allocation graph
τ	The number of periods it takes for the operator to receive a modular unit after it has been ordered
γ	The cost of a unit of influent over a time period
σ	The interest rate for the planning horizon
$\mathcal{A}$	The arcs in the skid allocation graph
$\mathcal{A}^{\text{new}}$	Arcs that allow for the transportation of new processing skids
$\mathcal{A}^{\text{pre}}$	Arcs that allow for the transportation of preexisting processing skids
$\mathrm{F}_{\text{node}}^{\text{path}}$	Function that returns the set of all ancestral nodes who share the same facility
$\mathrm{F}_{\text{node}}^{\text{new}}$	Function that returns a set of arcs (i, j), such that i is an ancestral of k that is at least τ periods older
$\mathrm{F}_{\text{forecast}}^{\text{node}}$	Function that returns the set of nodes that belong to a forecast
$\mathrm{F}_{\text{node}}^{\text{period}}$	Function that returns the period a node belongs to
$\mathrm{F}_{\text{node}}^{\text{kids}}$	Function that returns the first generation of descendants of a node in the skid allocation graph

$\mathrm{F}_{\mathrm{a}}^{\mathrm{desc.}}$	Function that returns the set of nodes that are descendants of a node in the skid allocation graph and belong to a specific forecast and period
$\mathrm{F}_{\mathrm{b}}^{\mathrm{desc.}}$	Function that returns the set of nodes that belong to a specific preexisting skid, forecast and period
$\mathrm{F}_{\mathrm{c}}^{\mathrm{desc.}}$	Function that returns the set of nodes that are descendants of a node in the skid allocation graph and belong to the same facility as the node and to a specific forecast
$\mathrm{P}_{\mathrm{new}}^{\mathrm{cap.}}$	The process capacity of a new skid with size/technology
$\mathrm{P}_{\mathrm{pre}}^{\mathrm{cap.}}$	The process capacity of a preexisting modular unit
$\mathrm{P}_{\mathrm{node}}^{\mathrm{influent}}$	The maximum influent flow-rate into node
$\mathrm{P}_{\mathrm{node}}^{\mathrm{prob\_path}}$	The probability of node given its ancestors
$\mathrm{P}_{\mathrm{node}}^{\mathrm{PW}}$	The present worth value of money at a time period for a node
$\mathrm{P}_{\mathrm{new}}^{\mathrm{cost}}$	The cost of a new processing skid of size and technology
$\mathrm{P}_{\mathrm{new}}^{\mathrm{OP}}$	The cost operating cost of a new processing skid of size and technology
$\mathrm{P}_{\mathrm{pre}}^{\mathrm{OP}}$	The cost operating cost of a preexisting processing skid
$\mathrm{P}_{\mathrm{trans}}^{\mathrm{cost}}$	The cost to transport a between facilities
$\mathrm{P}_{\mathrm{Cost}}^{\mathrm{HC}}$	The cost of waiting till a later period in the planning horizon to process the influent
$y_{i,j,k}^{\mathrm{new}}$	A binary variable that is equal to 1 if a new processing skid is initiated at node i of size and technology j and is currently located at node k, else it is equal to 0; such that $(i,j,k) \in \mathcal{A}^{\mathrm{new}}$
$y_{i,j,k,p}^{\mathrm{t\_new}}$	A binary variable that is equal to 1 if a new processing skid is initiated at node i of size and technology j and is currently located at node k and will be transported to node p in the next period, else it is equal to 0; such that $(i,j,k) \in \mathcal{A}^{\mathrm{new}}$ and $p \in \mathrm{F}_{\mathrm{node}}^{\mathrm{kids}}(k)$

$z^{\text{new}}_{i,j,k}$	A binary variable that is equal to 1 if a new processing skid is initiated at node i of size and technology j and is currently located at node k is operating, else it is equal to 0; such that $(i, j, k) \in \mathcal{A}^{\text{new}}$
$y^{\text{pre}}_{i,j}$	A binary variable that is equal to 1 if a preexisting processing skid i is currently located at node k, else it is equal to 0; such that $(i, k) \in \mathcal{A}^{\text{pre}}$
$y^{\text{t\_pre}}_{i,k,p}$	A binary variable that is equal to 1 if a preexisting processing skid i is currently located at node k and will be transported to node p in the next period, else it is equal to 0; such that $(i, k) \in \mathcal{A}^{\text{pre}}$ and $p \in \mathrm{F}^{\text{kids}}_{\text{node}}(k)$
$z^{\text{pre}}_{i,j}$	A binary variable that is equal to 1 if a preexisting processing skid i is currently located at node k is operating, else it is equal to 0; such that $(i, k) \in \mathcal{A}^{\text{pre}}$
x^{cap}_{k}	A non-negative variable that represents the capacity of all the modular units located at node $k \in \mathcal{N}$
$x^{\text{sch.}}_{i,k}$	A non-negative variable that is greater than 0 if the influent scheduled to be processed at node i is processed at node k, else it is equal to 0; such that $i, k \in \mathcal{N}$
x^{lost}_{k}	A non-negative variable that is the amount of influent for node, $k \in \mathcal{N}$, that was unprocessed in the time horizon and had to be disposed of
J_1	The cost for purchasing additional skids
J_2	The cost of operating all purchased skids
J_3	The cost of operating all preexisting skids
J_4	The cost to transport purchased skids between facilities
J_5	The cost to transport preexisting skids between facilities
J_6	The cost of holding off until a later period in the planning horizon to process the influent
J_7	The cost of not being able to process the influent within the planning horizon

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE | Page

1. INTRODUCTION AND LITERATURE REVIEW

It is projected, that by the year 2040 the amount of natural gas produced from shale formations will nearly double and account for almost two-thirds of all types of natural gas produced in the United States [1, 2]. It should also be noted that natural gas wells in shale formations typically have steep production decline curves in the first few years of operation that tail off as the wells production life increases [3]. This means after the first few years that a facility is in operation it is grossly oversized and therefore under utilized for the amount of shale gas it needs to process - assuming that the capacity of the facility is fixed and that only one well feeds the facility or multiple wells feed the facility but were completed at roughly the same. We postulate that if the facility was built with modular processing units that this excess processing capacity could be utilized in other regions that the operator is developing, which in turn would reduce the capital the operator of the exploration and production company would have to allocate to the new facility.

To date, the literature concerning multi-facility capacity expansion and contraction under uncertainty is quite sparse for both oilfield and general infrastructure planning. With regards to oilfield infrastructure planning, Aseeri, and Gorman, and Bagajewicz look at the optimal development of the offshore infrastructure under uncertainty [4]. Cafaro and Grossmann take a holistic view of oilfield planning and look at not only the size of gas processing plants but also the size of trunklines that connect the plants with the midstream lines, the power of gas compressors [5]. Cafaro and Grossmann only allow for the capacity of gas processing plants, gas compressors, and pipelines to expand and not contract in the planning horizon. Wang, Liang, Zheng, Lei, Yuan, and Zhang illustrate the benefits of shutting down and establishing new processing facilities for upstream pipeline networks so that the overall operational cost of the network is lowered [6].

Most of the traditional research, with regards to the operational research community, primarily focuses on capacity expansion problems and which do not allow the opportunity for the reallocation of excess capacity. For instance, Luss presents a survey of the literature concerning capacity expansion for deterministic and stochastic problems [7]. Bean, Higle, and Smith look at the case of capacity expansion for the infinite planning horizon under stochastic demands for unspecific processing equipment [8]. Neither one addresses the ability for the facilities to be comprised of modular units that can be reallocated between facilities to meet changes in demand over a planning horizon.

With that said, Naraharisetti and Karimi looked at allowing chemical plants to be multi-purpose as well as multi-production so that processing capacity could be transferred from the creation of one product to the creation of another product [9]. Jena, Cordeau, and Gendron look at capacity expansion and contraction through the use of a cost matrix that quantifies the cost from transitioning from one capacity to another [10]. As of late a few researchers have been looking at capacity expansion and contraction problems through the use of modular processing units for multi-facility problems. For instance, Melo, Nickle, and Da Gama did take the leap and allowed the facilities to be comprised of modular processing units that can be reallocated between facilities to meet changes in a deterministic demand [11]. Their algebraic model was formulated in such a way, that the capacity of every facility in every time period must be larger than the production demand for the corresponding time period. To the authors there has been no work concerning multi-stage stochastic capacity expansion and contraction problems that allows the postponement of processing if the capacity of the facility is insufficient to process all of the demand.

The purpose of this work is to generate a systematic methodology that can enable decision makers for exploration and production companies operating in shale formations to make better infrastructure planning decisions regarding there processing facilities. To accomplish this we have developed a novel approach, through the use of multi-stage stochastic programming, that allows each facility to be comprised of multiple modular transportable processing skids. The processing skids can be transported between facilities and purchased from manufacturers at discrete points within a finite planning horizon. We have also developed a novel recourse function that allows the decision maker to quantify the recourse actions that must be undertaken when the processing capacity of the facility is insufficient to process the influent. This in turn allows for a more flexible design of the facility network.

The remainder of this paper is structured as follows. First in Section 2 we present the necessary background on multi-stage stochastic programming. Then we transition into Section 3 that describes our motivating example in more detail. This is followed by Section 4 that states our problem statement. Then we give short explanation as to how the sizing of equipment in Section 5. After that we transition into Section 6, which describes the superstructure and novelties of our methodology. In Section 7, we describe how we mapped the superstructure into an algebraic model as well as state the constraints and objective functions. In Section 8 we describe our results and highlight the benefits of utilizing our methodology. Finally, we conclude with our closing remarks in Section 9.

2. BACKGROUND ON STOCHASTIC PROGRAMMING

Stochastic programming is a subset of mathematical programming and allows for the inclusion and quantification of uncertainties into optimization problems. Practically, stochastic programming is a methodology that allows a decision maker to make the optimal decision or set of decisions given a set of uncertain parameters whose exact values are not known a priori. These uncertainties are mapped into the optimization problem through the use of an expected value function or functions. If there is more than one expected value function, $\mathbb{E}$ in the stochastic program, it implies that information regarding the uncertainties is revealed to the decision maker at discrete points within a discretized time horizon. Each point or set of points within the discretized time horizon where the decision maker has the ability to make decisions regarding the uncertainties is referred to as a stage.

Stochastic programming problems can be classified by the number of stages that are incorporated within them: (i) single-stage, if only one stage is included in the problem; (ii) two-stage, if two stages are included in the problem; and (iii) multi-stage, if more than two stages are included in the problem. In single-stage stochastic programming the decision maker only has the ability to make "here-and-now" decisions, $x_1 \in \mathbb{X}_1$, which are decisions made in the first stage before the uncertainty, ξ_2, is revealed. An example of a single-stage stochastic program is:

$$\min_{x_1 \in \mathbb{X}_1} \left\{ f_1(x_1) + \mathbb{E}\left[f_2(\xi_2) \right] \right\}.$$

It should be noted that f_1 and f_2 quantify the cost of the first stage decision, x_1, before and after the uncertainty, ξ_2, is revealed respectively.

In two-stage stochastic programming the decision maker has the ability to make "here-and-now" decisions in the first stage as well as recourse actions, x_2, in the second stage. The recourse actions are decisions that are taken after the uncertainty has revealed itself and act as a corrective action to the first stage decision. It should be noted that in the literature, recourse actions are commonly referred to as "wait-and-see" decisions. An example of a two-stage stochastic program is:

$$\min_{x_1 \in \mathbb{X}_1} \left\{ f_1(x_1) + \mathbb{E}\left[\min_{x_2 \in \mathcal{X}_2} f_2(x_2, \xi_2) \right] \right\},$$

where $\mathcal{X}_2 := (\mathbb{X}_2 \mid x_1, \xi_2)$ is the feasible space of the recourse action and ensures that the recourse action is taken after the uncertainty has revealed itself. Depending on the formulation of the problem, the second stage decision can encompass multiple time periods. This type of problem is referred to as a two-stage multi-period stochastic program [12, 13].

In multi-stage stochastic programming the decision maker has the ability to make a set of S decisions, one set for each stage. The structure of the decision-making scheme, assuming the process is Markovian, can be seen below:

$$\begin{gathered}
\text{make the "here-and-now" decision, } x_1 \in \mathbb{X}_1 \\
\text{observe the uncertainty, } \xi_2 \\
\text{make the second stage decision, } x_2 \in (\mathbb{X}_2 \mid x_1, \xi_2) \\
\text{observe the uncertainty, } \xi_3 \\
\text{make the third stage decision, } x_3 \in (\mathbb{X}_3 \mid x_2, \xi_3, \xi_2) \\
\vdots \\
\text{observe the uncertainty, } \xi_S \\
\text{make the final stage decision, } x_S \in (\mathbb{X}_S \mid x_{S-1}, \xi_S, \xi_{S-1}).
\end{gathered}$$

.

It should be noted that the feasible solution region, $\mathbb{X}_s$, of every corrective action includes yet is not limited to non-anticipative constraints. Non-anticipative constraints ensure that the decision, x_s, for any stage must be made before the uncertainty in the following stage reveals itself. The mathematical form of the non-anticipative constraints, assuming the process is Markovian, can be written as: $(\mathbb{X}_s \mid x_{s-1}, \xi_s, \xi_{s-1}) \equiv (\mathbb{X}_s \mid A_s(\xi_s)x_{s-1}(\xi_{s-1}) + B_s x_s(\xi_s) = b_s(\xi_s))$. Here the matrices $A_s(\xi_s)$ and B_s and the vector $b_s(\xi_s)$ construct the non-anticipative constraints. The nested multi-stage stochastic programming representation of the above decisions scheme is as follows:

$$\min_{x_1 \in \mathbb{X}_1} \left\{ f_1(x_1) + \mathbb{E}\left[\min_{x_2 \in \mathcal{X}_2} f_2(x_2, \xi_2) + \mathbb{E}\left[\min_{x_3 \in \mathcal{X}_3} f_3(x_3, \xi_3) + \cdots + \mathbb{E}\left[\min_{x_S \in \mathcal{X}_S} f_S(x_S, \xi_S) \right]\right]\right]\right\},$$

where the solution space of the recourse actions is $\mathcal{X}_s := (\mathbb{X}_s \mid x_{s-1}, \xi_s, \xi_{s-1})$, assuming a Markovian structure. If the Markovian assumption is removed, and the decision made at stage s depends on all previous decisions then the solution space of the recourse actions can be restated as $\mathcal{X}_s := (\mathbb{X}_s \mid x_{s-1}, \ldots, x_1, \xi_s, \ldots, \xi_2) \equiv (\mathbb{X}_s \mid A_s(\xi_2)x_1 + \cdots + A_s(\xi_s)x_{s-1}(\xi_{s-1}) + B_s x_s(\xi_s) = b_s(\xi_s))$.

Stochastic programming problems are often converted to their deterministic equivalent for the sake of computation considerations. An example of the deterministic equivalent of a two-stage stochastic program is given as:

$$\begin{aligned} \min_{x_1,\, x_2^r} \;& f_1(x_1) + \sum_{r \in \mathcal{R}} p^r \cdot f_2^r(x_2^r) \\ \text{s.t.} \;& x_1 \in \mathbb{X}_1 \\ & x_2^r \in (\mathbb{X}_2 \mid x_1, r) \, \forall \, r \in \mathcal{R}, \end{aligned}$$

where $\mathcal{R}$ is a set of scenarios mapped from the uncertainties that reveal themselves in the second stage, p^r is the probability of the scenario $r \in \mathcal{R}$ transpiring, and $(\mathbb{X}_2 \mid x_1, r)$ is the feasible solution space for x_2^r and is equivalent to $(\mathbb{X}_2 \mid A_2^r x_1 + B x_2^r = b_2^r)$ [14].

The concept of non-anticipative constraints can be illustrated through the use of scenario graphs. Figure 2.1 represent two different types graphs that describe this principle: explicit and implicit scenario graphs. The nodes in each graph represents the decisions that can be made by the decision maker. The directed arcs, (i, j), which connect the nodes in each graph map the uncertainties and their associated probabilities, such that there exists a directed arc (i, j) from node i to node j in the graph.

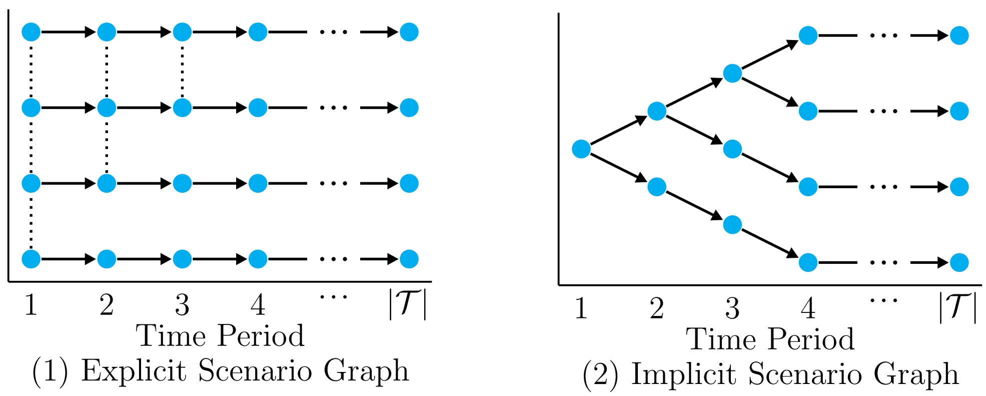


Figure 2.1: Scenario Graphs

Figure 2.1.1 is an explicit scenario graph with redundant nodes/variables that are bound by explicit non-anticipative constraints. The non-anticipative constraints are illustrated by the black dotted non-directed arcs and ensure that all nodes that are connected share the same decisions. Figure 2.1.2 is a graph with implicit non-anticipative constraints. The redundant nodes that are linked by the explicit non-anticipative constraints in Fig. 2.1.1 were removed and enforced implicitly by the structure of the graph [15]. It should be noted that by removing the excess nodes and explicit non-anticipative constraints the size of the stochastic programming problem is reduced; thereby, decreasing the computational time needed to solve the problem.

3. MOTIVATING EXAMPLE

Without loss of generality, in this section we describe a shale natural gas field where the operator of the exploration and production company is the decision maker, which can be seen in Fig. 3.1.

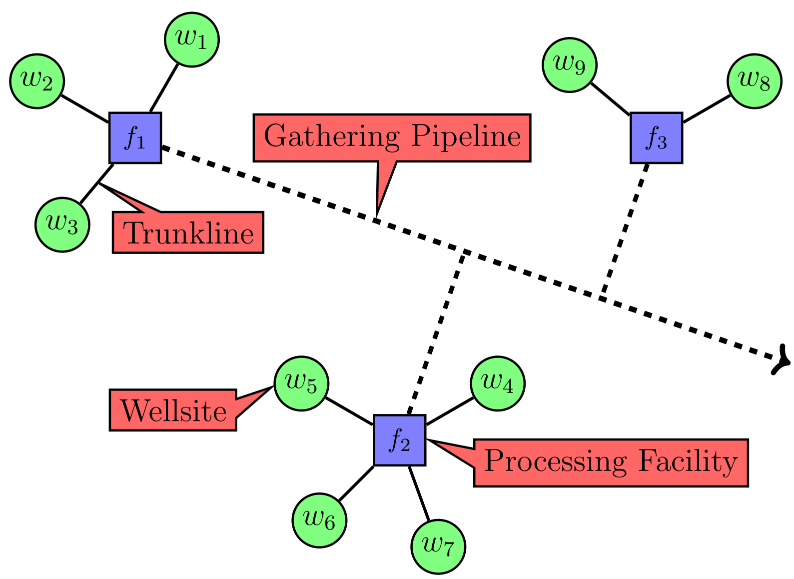


Figure 3.1: Hypothetical Field Under Development

In Fig. 3.1 there is as set of completed and proposed wells, $w \in \mathcal{W}$. The completed wells, are wells that have been drilled and are producing natural gas. The proposed wells, are wells that have been projected by the operator to be drilled at a later point in the planning horizon. The trunklines transport the unprocessed natural gas from the wellsites to the processing plants. The processing plants, $f \in \mathcal{L}$, which are also referred to as facilities within industry, process the effluent from the wells so that they can be transported to the midstream company through the gathering pipeline.

The following is a list of assumptions about the development of the aforementioned natural gas field:

1. The planning horizon can be discretized into a set of periods, $\mathcal{T}$;

2. The influents composition, pressure, and temperature into each facility remains constant for the planning horizon;

3. The influent flow-rate into each facility is assumed to be constant in each period in the discretized time horizon, this can be seen in Fig. 3.2 [5];

4. The uncertain production forecasts for each facility can be mapped into an implicit scenario graph.

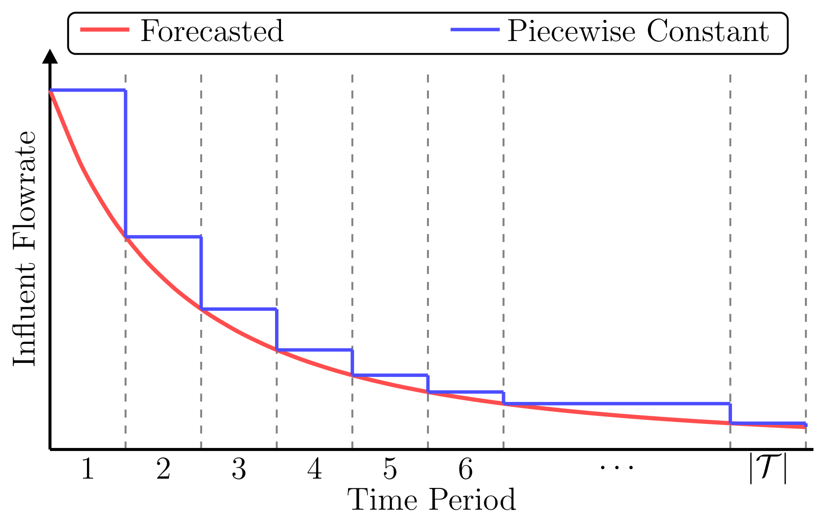


Figure 3.2: Conversion of Influent Flow-rate

Instead of using permanent facilities with fixed capacities, which is the traditional way that upstream facilities are built, the methodology presented allows the facilities to be comprised of multiple transportable modular skids. These modular skids can be transported via truck between facilities at every time discretization in the planning horizon. Moreover, the operator has the ability to generate a set of production forecasts, which simulate the uncertain nature of the influent into each facility. The production forecasts are mapped into an implicit scenario graph. The edges in the implicit scenario graph, map to the uncertainties and have an associated probability. Each node in the graph contains a collection of values. The values contained in the nodes represent the capacity each facility must have to ensure all of its influent is processed or measured in the corresponding time period. It should be noted that depending on the structure of the production forecasts, the implicit scenario graph will either map to a two-stage multi-period or a multi-stage stochastic program. For instance, if there is only one node in the implicit scenario graph with an out-degree greater than one it will map into a two-stage multi-period stochastic program. On the other hand, if there is more than one node with an out-degree greater than one, the implicit scenario graph will map into a multi-stage stochastic program.

There are two different sets of skids that can be utilized in the aforementioned network: preexisting skids and purchased skids. The set of preexisting skids are owned by the operator and can be utilized at any point within the planning horizon. It should be noted that it is possible in to not have any preexisting skids. The skids that are purchased by the operator must be chosen from a discrete set of processing capacities and technologies, $\mathcal{C}$. By allowing the skids to be transported between facilities and purchased at a later point in the planning horizon, the operator has increased his or her flexibility to combat the uncertainty that comes with developing facilities for an oilfield.

The following is a list of assumptions regarding the purchasing and reallocation of modular skids as well as the recourse function:

1. If the total processing capacity of a facility is less than the influent flow-rate into that facility for a time period in a forecast, then the excess influent must be processed in a later time period if there is excess capacity at the facility or it is assumed to be lost at the final stage;

2. Skids can be purchased from manufacturers or suppliers at all stages in the planning horizon except the final stage;

3. There is a lead time, τ, from when the operator orders a skid and when he receives it – it is not necessary that this time is greater than zero;

4. The operator must purchase skids from a discrete set of processing capacities and technologies;

5. Skids can be transported between facilities at a cost;

6. A maximum of $|\mathcal{C}|$ processing skids cab be purchased for at one facility for any given time period in a particular forecast, such that the capacity and technology of the skids that are purchased are all different and exist in $\mathcal{C}$;

7. A skid can only be located at one facility for any given time period in a particular forecast.

4. PROBLEM STATEMENT

Since the modular processing skids are mobile and additional skids can be purchased throughout the planning horizon, the operator has many decisions to make over the course of the planning horizon. These decisions can be broken down into three sets: first period decisions, intermediate period decisions, and final period decisions. As previously stated, depending on how the operator receives the information, the implicit scenario graph will either map to a two-stage multi-period or multi-stage stochastic program.

For the first period decisions, the operator has to decide: (a.i) how to allocate the pre-existing skids; (a.ii) should additional skids be purchased and if so what should be there respective capacity and technology; (a.iii) which skids should be operating during every time period for every forecast; and (a.iv) should the flow rate of the influent to each facility be reduced or halted when the capacity of the facility is insufficient to process the influent. The intermediate period decisions include all types of decisions made in the first period along with the need to determine if the excess influent from the previous time periods is processed at the current time period. The decision made at the final period only includes the ability to dispose of the excess unprocessed natural gas.

The above decisions were reformulated into a stochastic mixed integer linear programming problem that can be solved in a rolling horizon manner. This allows for: (b.i) the ability to purchase additional modular processing units; (b.ii) relocate existing skids; (b.iii) the ability to turn off a skid; and (b.iv) the ability to process the influent at a later time if there is excess capacity. The objective is to minimize the cost associated with the aforementioned decisions, which in turn will allow the operator to make the best "here-and-now" decisions so that his or her cost is minimized given the uncertain production forecasts.

5. EQUIPMENT SIZING

As stated in the motivating example, the size of the skids that comprise the facilities are chosen from a discrete set. The decision to force the capacity of the modular skids to come from a discrete set as opposed to belonging in a continuous range is twofold. Firstly, it ensures that there is not a discontinuous concave function that represents the cost of the processing unit in the objective function. Secondly, and more importantly, oilfield equipment is typically manufactured in discrete sizes. Therefore, if the size a of processing skid was allowed to be in a continuous range, then the solution of the problem could be sub-optimal. This is due to the fact that there would have to be posterior rounding to the discrete sizes that are available from suppliers [2].

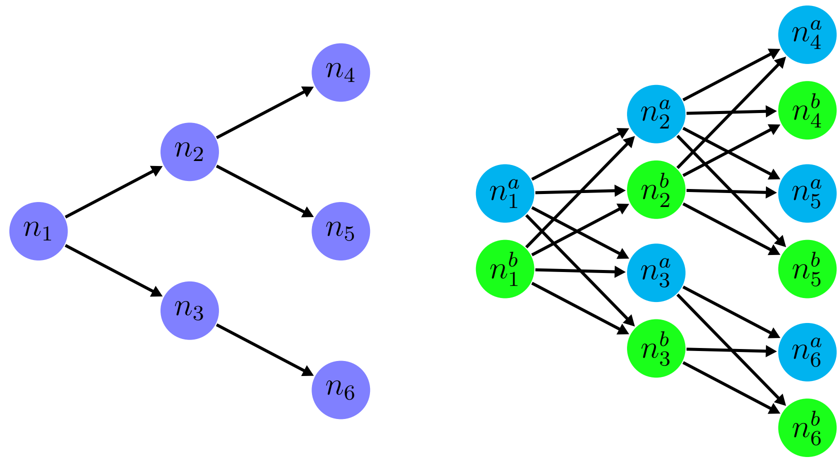
6. SUPERSTRUCTURE

We have developed a novel superstructure that can be formulated as a two-stage multi-period or a multi-stage stochastic program depending on the structure of the uncertain production forecasts. This is accomplished by utilizing implict non-anticipative constraints, which have been mapped into the superstructure through the use of the implicit scenario graph. The superstructure is comprised of two sets of decision types: (i) purchasing and allocation actions, as well as (ii) recourse actions. The two sets of decision types are joined together by accounting for the capacity of each facility at every time period in the planning horizon. To the authors knowledge this approach has not been undertaken before.

6.1 Purchasing and Allocation Actions

We have developed a decision-making scheme that allows for the purchasing of additional modular processing units for every facility in every forecast at every time period. This scheme also allows the ability to relocate any preexisting processing skids or purchased skids between facilities at discretized points in the planning horizon. This is accomplished by mapping the implicit scenario graph into an allocation graph.

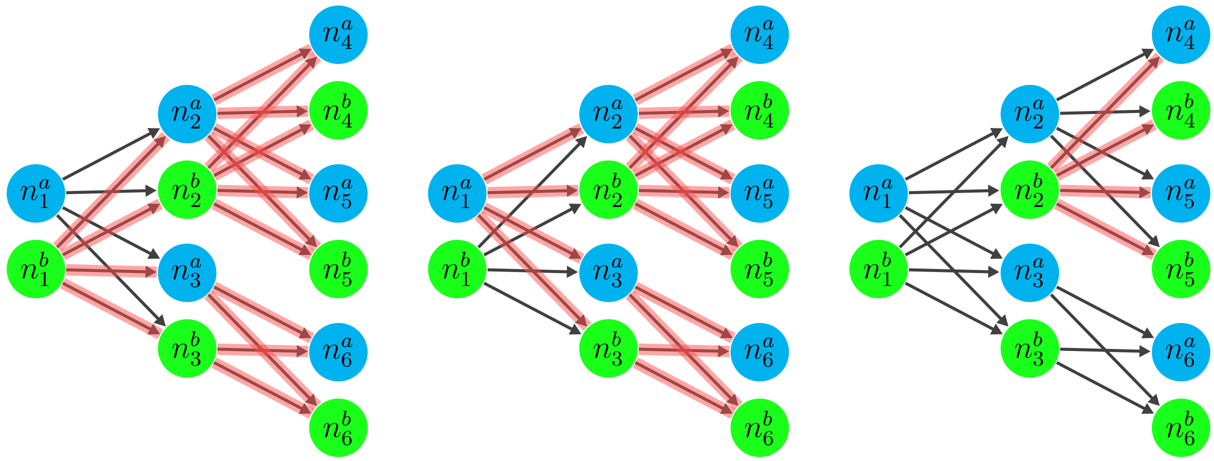
For sake of an example and without loss of generality, we assume the operator is trying to develop a field given the following parameters: (c.i) a planning horizon of three periods; (c.ii) a field that has two facilities, $\{a, b\}$; and (c.iii) three production forecasts. Given these parameters, we assume that the operator can build an implicit scenario graph, which can be seen in Fig. 6.1.1.



(1) Implicit Scenario Graph (2) Skid Allocation Graph

Figure 6.1: Scenario and Allocation Graph

For this example, the forecasts and periods are arranged as follows: (d.i.a) forecast 1 includes the set of nodes, $\{n_1, n_2, n_4\}$; (d.i.b) forecast 2 includes the set of nodes, $\{n_1, n_2, n_5\}$; and (d.i.c) forecast 3 includes the set of nodes, $\{n_1, n_3, n_6\}$. (d.ii.a) the first period decisions includes the set of nodes; $\{n_1\}$ (d.ii.b) the second period decisions includes the set of nodes, $\{n_2, n_3\}$; and (d.ii.c) the third period decisions includes the set of nodes, $\{n_4, n_5, n_6\}$. Given the implicit scenario graph, we are able to map it into a skid allocation graph as seen in Fig. 6.1.2 by utilizing the aforementioned assumptions regarding the reallocation of modular skids. There are now nodes that symbolize every facility and period for every forecast, such that the facilities exists in the set $\{a, b\}$. Given the skid allocation graph in Fig. 6.1.2, the basic idea of how preexisting and purchased processing skids can be reallocated in a natural gas field is illustrated in Fig. 6.2. Where the highlighted directed arcs in Fig. 6.2 allow a skid to be transported from node i to node j, such that there is a directed arc (i, j) that connect them.



(1) Preexisting Skid, 1, Located at Facility b (2) Purchased Skid at n_1^a (3) Purchased Skid at n_2^b

Figure 6.2: Graphs for Purchasing and Allocations Actions

It should be noted, as previously stated in the assumptions that a processing skid can only be located at one facility for any given time period in a forecast. For example in Fig. 6.2.1, the preexisting skid that was originally located at node, n_1^b, can be transported to either n_2^a or n_2^b and either n_3^a or n_3^b. This is due to the fact that n_2 and n_3 do not belong to the same forecast. This same principle applies for purchased skids. For instance, in Fig. 6.2.3, the skid purchased at node n_2^b can be transported to either n_4^a or n_4^b and either n_5^a or n_5^b.

6.2 Recourse Actions

We have also developed a decision-making scheme that enables the processing of the influent for every facility in every forecast to be postponed. This in turn, allows the operator to process excess unprocessed influent at a later point in the planning horizon due to insufficient processing capacity. The manner in which we formulated the decision-making scheme also implicitly enforces non-anticipative constraints.

Without loss of generality, Fig. 6.3 is a graphical example of the set of recourse actions that must be taken given a hypothetical policy for a given facility and forecast due to insufficient processing capacity.

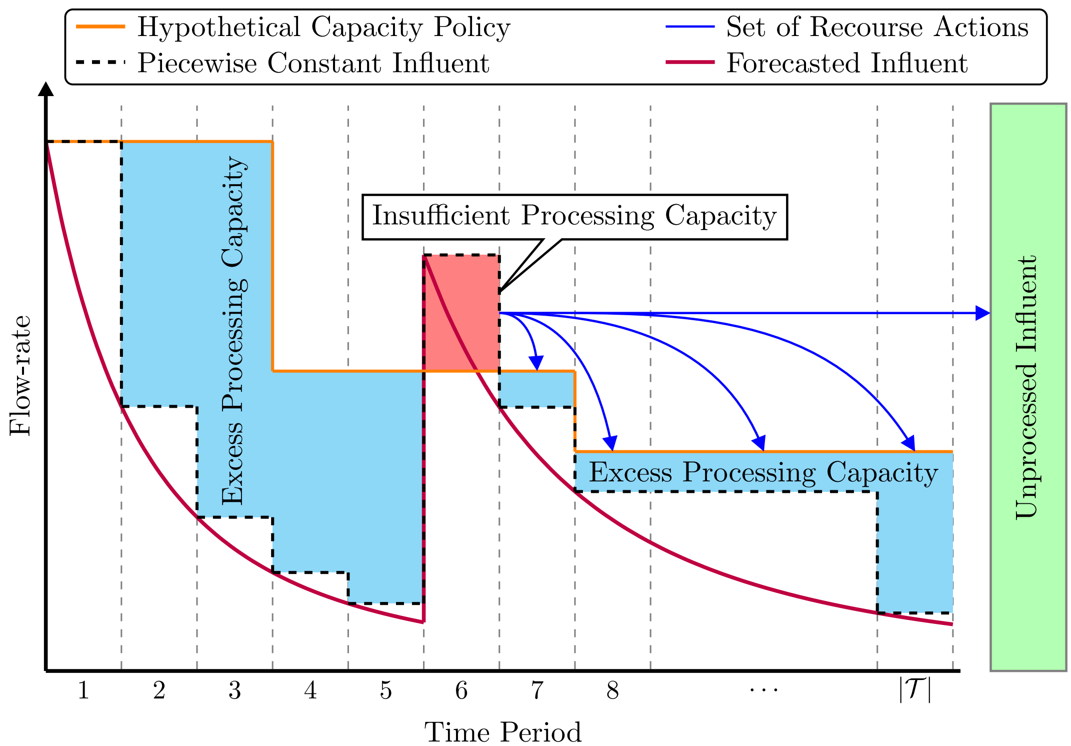


Figure 6.3: Hypothetical Policy for a Facility and Forecast

From inspection of Fig. 6.3 it is clear that processing capacity of the facility is insufficient to process all of the influent. Therefore, the operator has to reduce the flow rate of the effluent of the wells that feed the facility under consideration. This ensures the influent is not larger than the total processing capacity of the facility. Moreover, the operator has the ability to process the influent at a later point in the planning horizon when there is excess processing capacity or it could remain unprocessed. If excess effluent remains unprocessed the gas is assumed to be flared.

7. ALGEBRAIC MODEL

The algebraic model of our superstructure contains both binary and continuous variables. The binary variables allow for the purchasing of additional processing units, the ability to turn off a purchased or preexisting skid for a time period, and the ability to move the skids between facilities. The continuous variables allow for the processing capacity of facilities, the amount of gas each facility processes, and the amount of gas that each facility is unable to process to be accounted for in all periods for all forecasts. These variables are then utilized to quantify the decisions outlined in the problem statement.

7.1 Arcs

There are two sets of arcs that act as the backbone for the indices of the variables in our algebraic model. The arcs in $\mathcal{A}^{\text{new}}$ represent the indices for the variables concerned with the purchased modular processing units. The arcs that exist in $\mathcal{A}^{\text{new}}$ are a subset of $\{(i,j,k) \in \mathcal{N} \times \mathcal{C} \times \mathcal{N}\}$; such that $\mathcal{N}$ is the set of nodes in the skid allocation graph and the nodes i and k are connected in the skid allocation graph or they are identical. These arcs allow for tracking the nodes that a skid purchased at node i and with a capacity and technology j reaches, k. The arcs that exist in $\mathcal{A}^{\text{pre}}$ represent the indices for the preexisting modular units. This set is a subset of $\{(i,j) \in \mathcal{B} \times \mathcal{N}\}$, where j is a node that the preexisting skid i can reach. Practically, this is an initial condition and ensures that skid i is located at the correct node at the first period.

7.2 Variables

7.2.1 Binary Variables

There are three different binary variables concerned with new modular processing skids. The binary variable $y^{\text{new}}_{i,j,k}$, such that (i,j,k) exists in $\mathcal{A}^{\text{new}}$, is equal to one if a new processing skid is initiated at node i of capacity and technology j and is currently located at node k, otherwise it is equal to zero. The variable $y^{\text{t\_new}}_{i,j,k,p}$, such that $(i,j,k) \in \mathcal{A}^{\text{new}}$ and $p \in \mathrm{F}^{\text{kids}}_{\text{node}}(k)$, is equal to one if a new processing skid is initiated at node i of capacity and technology j and is currently located at node k and will be transported to node p in the next period otherwise it is equal to zero. Here $\mathrm{F}_{\mathrm{node}}^{\mathrm{kids}}(\cdot)$ is a function that returns the first generation of descendants of a node in the skid allocation graph. The variable $z_{i,j,k}^{\mathrm{new}}$, such that $(i,j,k) \in \mathcal{A}^{\mathrm{new}}$, is equal to one if a new processing skid is initiated at node i of capacity and technology j and is currently located and operating at node k, otherwise it is equal to zero.

There are three different binary variables concerned with preexisting modular processing skids. The variable $y_{i,k}^{\mathrm{pre}}$, such that $(i,k) \in \mathcal{A}^{\mathrm{pre}}$, is equal to one if a preexisting processing skid i is currently located at node k, otherwise it is zero. The variable $y_{i,k,p}^{\mathrm{t\_pre}}$, such that $(i,k) \in \mathcal{A}^{\mathrm{pre}}$ and $p \in \mathrm{F}_{\mathrm{node}}^{\mathrm{kids}}(k)$, is equal to one if a preexisting processing skid i is currently located at node k and will be transported to node p in the next period, otherwise it is equal to zero. The variable $z_{i,j}^{\mathrm{pre}}$, such that $(i,k) \in \mathcal{A}^{\mathrm{pre}}$, is equal to one if a preexisting processing skid i is currently located and operating at node k, otherwise it is equal to zero.

7.2.2 Non-negative Variables

There are three different non-negative continuous variables utilized in the algebraic model. The variable x_{k}^{cap}, such that $k \in \mathcal{N}$, represents the combined capacity of all the preexisting and purchased skids located at node k. If the value of the variable $x_{i,k}^{\mathrm{sch.}}$, such that $i,k \in \mathcal{N}$ and i is equal to k or node k is reachable from node i in the skid allocation graph, is greater than zero the influent scheduled to be processed at node i is processed at node k. Practically, the variable $x_{i,k}^{\mathrm{sch.}}$ allows the influent that should be processed at the time period, facility, and forecast to which node i belongs, to be processed at a later time period that belongs to the same facility and forecast. The variable x_{k}^{lost}, such that $k \in \mathcal{N}$, is the amount of influent for node k that was unable to be processed in the time horizon and had to be disposed of.

7.3 Constraints

7.3.1 New Processing Skids

Equation 7.1 ensures that if a modular processing unit of capacity and technology, j, is purchased at node i, then it can only be located at one facility for every forecast that node i belongs too. On the other hand, if there is no skid of capacity and technology, j, purchased at node i, then none of node i's descendants in the skid allocation graph, $p \in \mathcal{N}$, can be active; such that $i, p \in \mathcal{N}$ and $j \in \mathcal{C}$.

$$y_{i,j,i}^{\text{new}} = \sum_{p \in \mathrm{F}_{\mathrm{a}}^{\text{desc.}}(i,f,t)} y_{i,j,p}^{\text{new}} \;\; \forall\, j \in \mathcal{C},\, f \in \mathcal{F},\, i \in \mathrm{F}_{\text{forecast}}^{\text{node}}(f),\, t \in \mathcal{T} \mid t > \mathrm{F}_{\text{node}}^{\text{period}}(i) \tag{7.1}$$

The function $\mathrm{F}_{\mathrm{a}}^{\text{desc.}}(i, f, t)$, such that $i \in \mathcal{N}$, $f \in \mathcal{F}$ and $t \in \mathcal{T}$, returns the set nodes that are descendants of node i and belong to forecast f and period t. The function $\mathrm{F}_{\text{forecast}}^{\text{node}}(f)$ returns the set of nodes that belong to the forecast f and the function $\mathrm{F}_{\text{node}}^{\text{period}}(i)$ returns the period of a node.

Equation 7.2 allows the transportation of skids to be tracked between facilities through the planning horizon, which in turn allows the transportation cost to be included in the objective function.

$$y_{i,j,k}^{\text{new}} + y_{i,j,p}^{\text{new}} \leq 1 + y_{i,j,k,p}^{\text{t\_new}} \;\; \forall\, (i,j,k) \in \mathcal{A}^{\text{new}},\, p \in \mathrm{F}_{\text{node}}^{\text{kids}}(k) \tag{7.2}$$

Equation 7.3 enforces an upper bound on the binary variable, $z_{i,j,k}^{\text{new}}$, such that $(i, j, k) \in \mathcal{A}^{\text{new}}$. This ensures that a new processing skid of capacity and technology j that originated at node i can only be operational at k if $y_{i,j,k}^{\text{new}}$ is equal to one.

$$y_{i,j,k}^{\text{new}} \geq z_{i,j,k}^{\text{new}} \;\; \forall\, (i,j,k) \in \mathcal{A}^{\text{new}} \tag{7.3}$$

7.3.2 Preexisting Processing Skids

Equation 7.4 ensures that a preexisting skid can only be located at a single facility for every time period in every forecast. Here, the function $\mathrm{F}_{\mathrm{b}}^{\mathrm{desc.}}(i, f, t)$, such that $i \in \mathcal{B}$, $f \in \mathcal{F}$ and $t \in \mathcal{T}$, returns the set of nodes that belong to the preexisting skid i, to the forecast f, and the period t.

$$\sum_{p \in \mathrm{F}_{\mathrm{b}}^{\mathrm{desc.}}(i,f,t)} y_{i,p}^{\mathrm{pre}} = 1 \ \ \forall \, i \in \mathcal{P},\, f \in \mathcal{F},\, t \in \mathcal{T} \tag{7.4}$$

Equation 7.5 is similar to Eq. 7.2, in that it allows the skids to be spatially tracked through the field. However, in this case, the skid that is being tracked is a preexisting skid.

$$y_{i,k}^{\mathrm{pre}} + y_{i,p}^{\mathrm{pre}} \leq 1 + y_{i,k,p}^{\mathrm{t\_pre}} \ \ \forall \, (i,k) \in \mathcal{A}^{\mathrm{pre}},\, p \in \mathrm{F}_{\mathrm{node}}^{\mathrm{kids}}(k) \tag{7.5}$$

Equation 7.6 is similar to Eq. 7.3; such that it allows a skid to be located at a facility but not operating. However, in this case, it is for preexisting skids not purchased skids.

$$y_{i,k}^{\mathrm{pre}} \geq z_{i,k}^{\mathrm{pre}} \ \ \forall \, (i,k) \in \mathcal{A}^{\mathrm{pre}} \tag{7.6}$$

7.3.3 Processing Capacity of Facilities

Equation 7.7 is an equality constraint that ensures the variable x_k^{cap}, such that $k \in \mathcal{N}$, is equal to the capacity of all the skids located at the node k. So when the binary variables, $z_{i,j,k}^{\mathrm{new}}$ and $z_{p,k}^{\mathrm{pre}}$ are equal to one, their respective processing capacities, $\mathrm{P}^{\mathrm{new}}(j)$ and $\mathrm{P}^{\mathrm{pre}}(p)$, such that $j \in \mathcal{C}$ and $p \in \mathcal{B}$, are included in the summation term.

$$x_k^{\mathrm{cap}} = \sum_{(i,j) \in \mathrm{F}_{\mathrm{node}}^{\mathrm{new}}(k,\tau)} \mathrm{P}^{\mathrm{new}}(j) \cdot z_{i,j,k}^{\mathrm{new}} + \sum_{(p,k) \in \mathcal{A}^{\mathrm{pre}}} \mathrm{P}^{\mathrm{pre}}(p) \cdot z_{p,k}^{\mathrm{pre}} \ \ \forall \, k \in \mathcal{N} \tag{7.7}$$

The function $\mathrm{F}_{\mathrm{node}}^{\mathrm{new}}(k, \tau)$ returns a set of arcs (i, j), such that i is an ancestral node of k that is at least τ periods older and $(i, j, k) \in \mathcal{A}^{\mathrm{new}}$.

7.3.4 Facility Utilization

Equation 7.8 allows the influent that should be processed at node i to be processed at a later time period or it is flared.

$$x_i^{\text{flared}} + \sum_{k \in \mathrm{F}_{\mathrm{c}}^{\text{desc.}}(i,f)} x_{i,k}^{\text{sch.}} = \mathrm{P}^{\text{influent}}(i) \;\; \forall\, f \in \mathcal{F},\, i \in \mathrm{F}_{\text{forecast}}^{\text{node}}(f) \tag{7.8}$$

Where $\mathrm{F}_{\mathrm{c}}^{\text{desc.}}(i,f)$, such that $i \in \mathcal{N}$ and $f \in \mathcal{F}$, is a function that returns the set of nodes that are descendants of node i, in the skid allocation graph, belong to the same facility as node i and belong to the forecast f. The parameter $\mathrm{P}^{\text{influent}}(i)$ is equal to the amount of influent forecasted for node i.

Equation 7.9 ensures that the processing capacity of node i, represented by the variable $x_i^{\text{cap.}}$, such that $i \in \mathcal{N}$, is greater than the amount of natural gas that node is supposed to process.

$$\sum_{k \in \mathrm{F}_{\text{node}}^{\text{path}}(i)} x_{i,k}^{\text{sch.}} \leq x_i^{\text{cap.}} \;\; \forall\, i \in \mathcal{N} \tag{7.9}$$

Where $\mathrm{F}_{\text{node}}^{\text{path}}(i)$, such that $i \in \mathcal{N}$, is a function that returns the set of all ancestral nodes of i who share the same facility and well as i.

7.4 Objective Functions

The following is a list of objective functions, which quantify the decisions that the operator needs to make as stated in the problem statement.

The objective function J_1 is the cost of purchasing additional skids over the planning horizon,

$$\mathrm{J}_1 = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{C}} \mathrm{P}_{\text{node}}^{\text{prob\_path}}(k) \cdot \mathrm{P}_{\text{node}}^{\text{PW}}(k) \cdot \mathrm{P}_{\text{new}}^{\text{cost}}(j) \cdot y_{i,j,k}^{\text{new}}, \tag{7.10a}$$

where $\mathrm{P}_{\text{node}}^{\text{PW}}(k)$, such that $k \in \mathcal{N}$, is the present worth of a decision made at node k based upon an interest rate. The parameter $\mathrm{P}_{\text{node}}^{\text{prob\_path}}(k)$, such that $k \in \mathcal{N}$, is the probability of node k occurring given its ancestral path. The parameter $\mathrm{P}_{\text{new}}^{\text{cost}}(j)$, such that $j \in \mathcal{C}$, is the cost of a new processing skid of capacity and technology j.

The objective function J_2 is the cost of operating all purchased skids,

$$\mathrm{J}_2 = \sum_{(i,j,k)\in\mathcal{A}^{\text{new}}} \mathrm{P}^{\text{prob\_path}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{PW}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{OP}}_{\text{new}}(j) \cdot z^{\text{new}}_{i,j,k}, \tag{7.10b}$$

where $\mathrm{P}^{\text{OP}}_{\text{new}}(j)$, such that $j \in \mathcal{C}$, is the operating cost of a new processing skid of capacity and technology j.

The objective function J_3 is the cost of operating all preexisting skids,

$$\mathrm{J}_3 = \sum_{(i,j)\in\mathcal{A}^{\text{pre}}} \mathrm{P}^{\text{prob\_path}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{PW}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{OP}}_{\text{pre}}(i) \cdot z^{\text{pre}}_{i,k}, \tag{7.10c}$$

where $\mathrm{P}^{\text{OP}}_{\text{pre}}(i)$, such that $i \in \mathcal{B}$, is the operating cost of a preexisting processing skid i.

The objective function J_4 is the cost to transport purchased skids between facilities,

$$\mathrm{J}_4 = \sum_{(i,j,k)\in\mathcal{A}^{\text{new}}} \sum_{p\in\mathrm{F}^{\text{kids}}_{\text{node}}(k)} \mathrm{P}^{\text{prob\_path}}_{\text{node}}(p) \cdot \mathrm{P}^{\text{PW}}_{\text{node}}(p) \cdot \mathrm{P}^{\text{cost}}_{\text{trans}}(k,p) \cdot z^{\text{t\_new}}_{i,j,k,p}, \tag{7.10d}$$

where $\mathrm{P}^{\text{cost}}_{\text{trans}}(k,p)$, such that $k,p \in \mathcal{N}$, is the cost to transport a new processing skid from node k to node p.

The objective function J_5 is the cost to transport preexisting skids between facilities,

$$\mathrm{J}_5 = \sum_{(i,k)\in\mathcal{A}^{\text{pre}}} \sum_{p\in\mathrm{F}^{\text{kids}}_{\text{node}}(k)} \mathrm{P}^{\text{prob\_path}}_{\text{node}}(p) \cdot \mathrm{P}^{\text{PW}}_{\text{node}}(p) \cdot \mathrm{P}^{\text{cost}}_{\text{trans}}(k,p) \cdot z^{\text{t\_pre}}_{i,k,p}. \tag{7.10e}$$

The objective function J_6 is the cost of holding off until a later period in the planning horizon to process the influent,

$$\mathrm{J}_6 = \gamma \cdot \sum_{i\in\mathcal{N}} \sum_{k\in\mathrm{F}^{\text{path}}_{\text{node}}(i)} \mathrm{P}^{\text{prob\_path}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{PW}}_{\text{node}}(k) \cdot \mathrm{P}^{\text{HC}}_{\text{node}}(i,k) \cdot x^{\text{sch.}}_{i,k}, \tag{7.10f}$$

where γ is the cost per unit of influent over a time period and $\mathrm{P}^{\text{HC}}_{\text{node}}(i,k)$, such that $i,k \in \mathcal{N}$

and k is reachable from i in the skid allocation graph or i is equal to k, is the cost to wait untill a later time period in the planning horizon to process the influent.

The objective function J_7 is the cost of not being able to process the influent within the planning horizon,

$$\mathrm{J}_7 = \gamma \cdot \sum_{i \in \mathcal{N}} \mathrm{P}_{\mathrm{node}}^{\mathrm{prob\_path}}(i) \cdot \mathrm{P}_{\mathrm{node}}^{\mathrm{PW}}(i) \cdot x_i^{\mathrm{lost}}. \tag{7.10g}$$

Given the algebraic constraints, Eqs. (1-9), and the objective functions, Eqs. (10a-10g), the formal definition of the stochastic program is as follows:

$$\begin{aligned} \min_{x,y,z} \quad & J_1 + J_2 + J_3 + J_4 + J_5 + J_6 + J_7 \\ \text{s.t.} \quad & \text{Equality \& Inequality Constraints, Eqs. } (1-9) \end{aligned}$$

8. RESULTS AND DISCUSSION

We utilized the implicit scenario graph, as seen in Fig. A.1, and the data in Table A.1 to create the structure of the algebraic model. The data for the influent flow-rates into each of the facilities in Table A.1 is hypothetical and was created for the sake of example, since industry data was unavailable. The length of each time period is assumed to be six months. The interest rate is assumed to be 6.25% per time period. The cost of natural gas is assumed to be 2.75 [\$/MSCF]; therefore, γ is approximated to $500,000$ [\$/MMSCF] per time period. The lead time to receive a skid after it has been ordered, τ, is one time period. There is only one type processing technology. The processing capacities of skids are $\{10,\, 20, \ldots,\, 130\}$ [MMSCFD], the highest capacity set to 130 because that is the maximum influent flow-rate any facility will see over the course of its operating horizon. The operating cost for each skid per time period is assumed to be 2.5% of the original purchase cost. The transportation cost to reallocate a skid from one facility to another is assumed to be 5% of the original purchase cost. The cost of new processing skids is based upon the sixth-tenth-factor-rule,

$$\text{cost\_of\_skid}_{\text{A}} = \text{cost\_of\_skid}_{\text{B}} \left(\frac{\text{size\_of\_skid}_{\text{A}}}{\text{size\_of\_skid}_{\text{B}}} \right)^{0.65},$$

here we assume that a skid of that has a capacity of 20 [MMSCFD] cost \$50,000 [16]. We assume there are two preexisting skids that are located at facility A: 'Preexisting skid 1' that it has a capacity of 50 [MMSCFD] and 'Preexisting skid 2' that it has a capacity of 70 [MMSCFD].

The algebraic model of the stochastic program was constructed with the Gurobi Python interface. The program was on solved on a machine with a 2.8 GHz Intel Core i7 processor and 16 GM of RAM utilizing Gurobi V8.0.0 [17].

8.1 Results

The aforementioned optimization problem was solved via Gurobi in 536.01 seconds with an optimality gap of 0.0000% and the cost of the optimal policy was found to be $223,929.00. Figure 8.1 is a graphical representation of the results that illustrates the location of the preexisting and hypothetical skids through the planning horizon for every forecast.

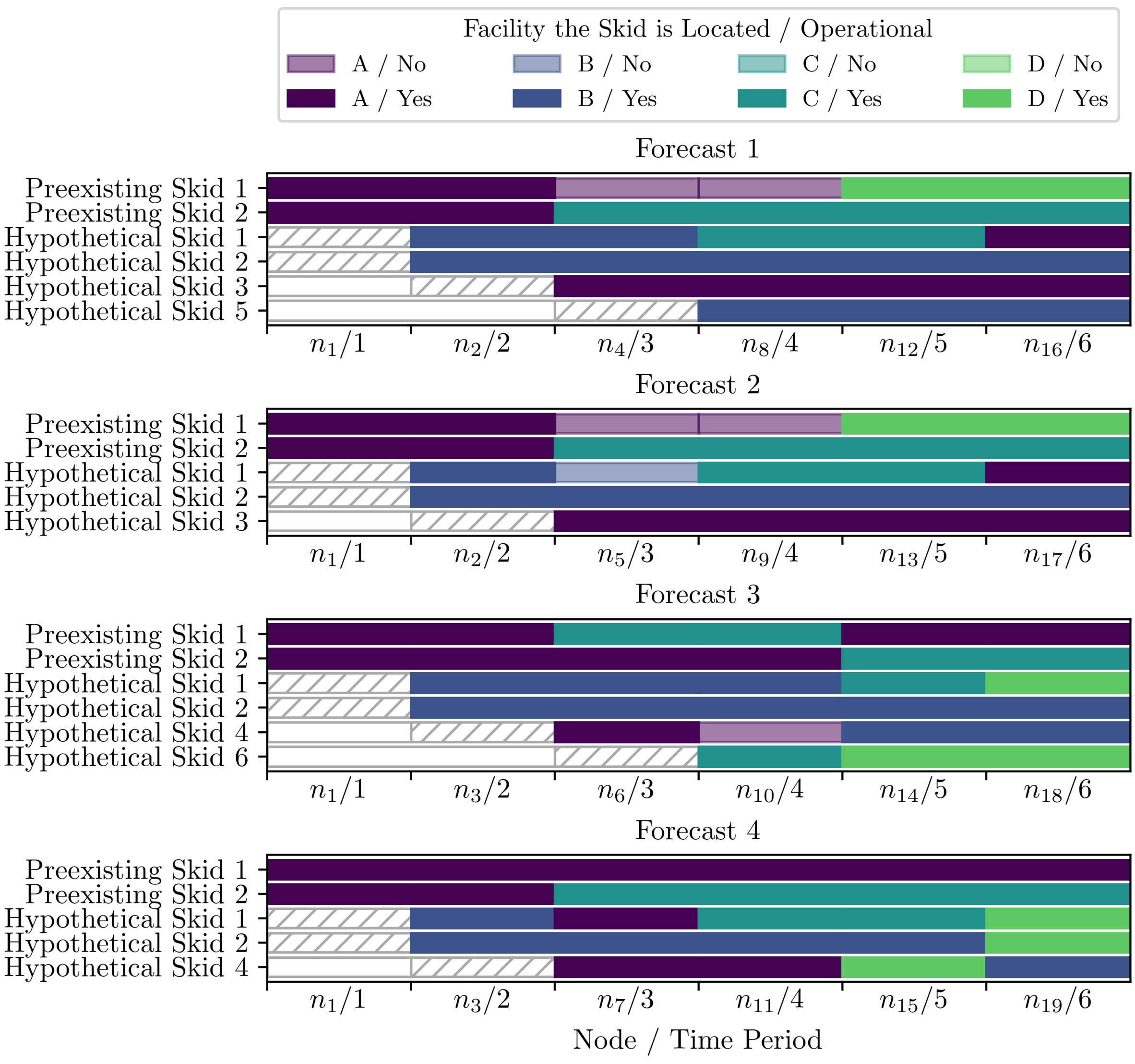


Figure 8.1: Ordering and Allocation Schedule for Different Forecasts

Table 8.1 illustrates the processing capacities [MMSCFD] for each of the hypothetical processing skids described in Fig. 8.1. For the "here-and-now" first stage decisions, the operator should purchase "Hypothetical Skid 1" and "Hypothetical Skid 2" as well as leave "Preexisting Skid 1" and "Preexisting Skid 2" at "Facility A". It should be noted, that the hatched regions in Fig. 8.1 represent the time period the processing skid was purchased from the supplier. It should also be noted in Fig. 8.1 the opaque regions indicate that a processing skid is located at a facility but is not operating.

Table 8.1: Capacities of Hypothetical Processing Skids

Processing Skid Name	Processing Capacity [MMSCFD]
Hypothetical Skid 1	10
Hypothetical Skid 2	40
Hypothetical Skid 3	120
Hypothetical Skid 4	20
Hypothetical Skid 5	20
Hypothetical Skid 6	40

Figure 8.2 is a graphical representation of the capacity of every facility for each forecast. The dashed blue lines represent the operating capacity of the facility and the solid red lines represent the piecewise constant influent flow-rate.

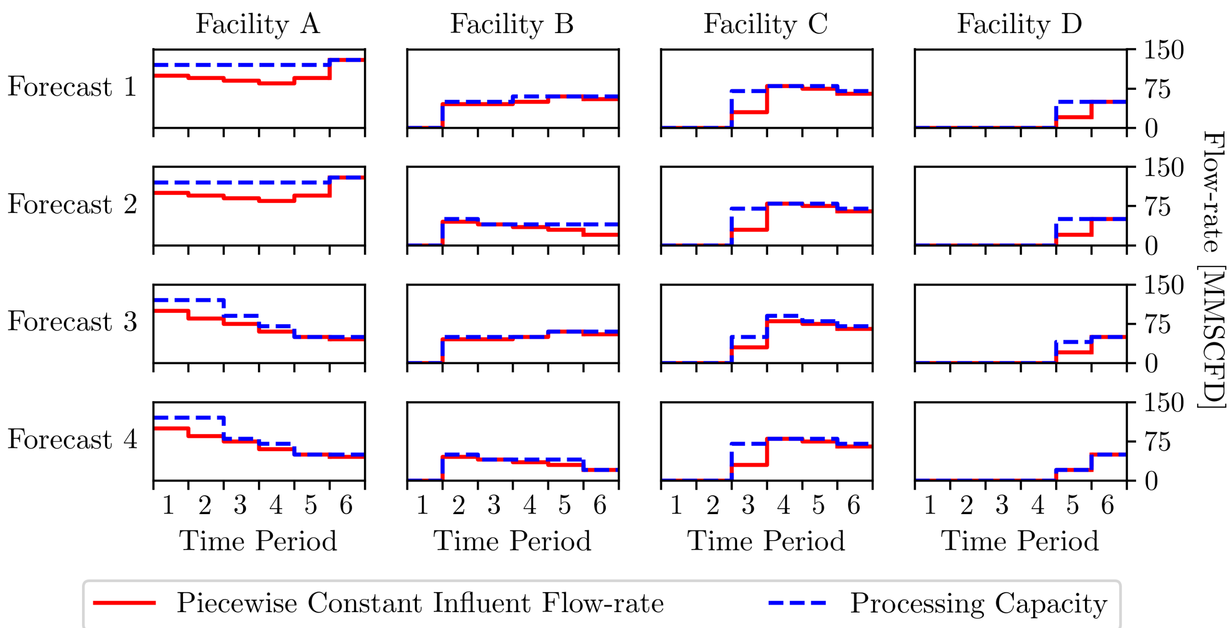


Figure 8.2: Processing Capacity of Facilities for Different Forecasts

8.2 Discussion

Due to the nature of the problem formulation, the maximum upper bound of our proposed methodology of utilizing flexible facilities that are comprised of transportable processing skids is the cost of utilizing permanent facilities, assuming that: (e.i) the cost of a permanent facility is the same as the cost of a mobile processing unit of the same capacity and technology; (e.ii) the capacities and technologies that can be utilized in the permanent facilities are identical to the ones that can be utilized for the transportable processing skids; and (e.iii) the capacity of each permanent facility is large enough that is able always process all of its influent for its useful life. This is clear from induction, since each facility is comprised of one skid that is never relocated to a different facility.

For our motivating example, we found that our proposed method methodology indicated that the operator should make the "here-and-now" decisions to immediately purchase "Hypothetical Skid 1" and "Hypothetical Skid 2", not relocate either of the preexisting skids, and process all of the influent for "Facility A". From inspection of Fig. 8.1, the location of these two skids, "Hypothetical Skid 1" and "Hypothetical Skid 2", and their operational status varied for each of the four different production forecasts. In addition, it is clear that the non-anticipative constraints ensured that decisions regarding these two skids, as well as, the other preexisting and hypothetical skids were made before the uncertainties revealed themselves. Also, with with respect to Fig. 8.2, it is clear the capacities of each of the facilities is always larger than the influent flow-rate to the facilities. However, if the market price of natural gas was lower it could be more financially beneficial to purchase lower capacity skids so that the cost of insufficient processing capacity would be offset by the savings in capital investments.

We ran random production forecasts as test cases to compare the cost of permanent facilities to the cost of flexible facilities, which can be seen in Fig. 8.3. The following assumptions were made regarding the generation of the random production forecasts: (f.i) the time horizon is four years; (f.ii) the time period length six months; (f.iii) the test cases were mapped into a two-stage multi-period stochastic program for the sake of computational time; (f.iv) the maximum influent that each facility would receive was uniformly selected between 50 and 200 for each test case; (f.v) each forecast within the test cases monotonically increased or decreased according to the distribution, Beta($\alpha = 1.5, \beta = 30$), and a scaling factor; (f.v.a) the scaling factor normalized the increase or decrease between between time periods based upon the maximum influent the facility could see; (f.v.b) a monotonically increasing forecast can switch to decreasing forecast with a probability of 0.25 at every time period; (f.vi) at least one processing plant started to receive influents in the first time period, the others started to receive influents uniformly between the first time period and the final time period; (f.vii) if a plant started to receive influents in the first time period two forecasts would be generated: one that is increasing and the other is decreasing; (f.vii.b) the probability.

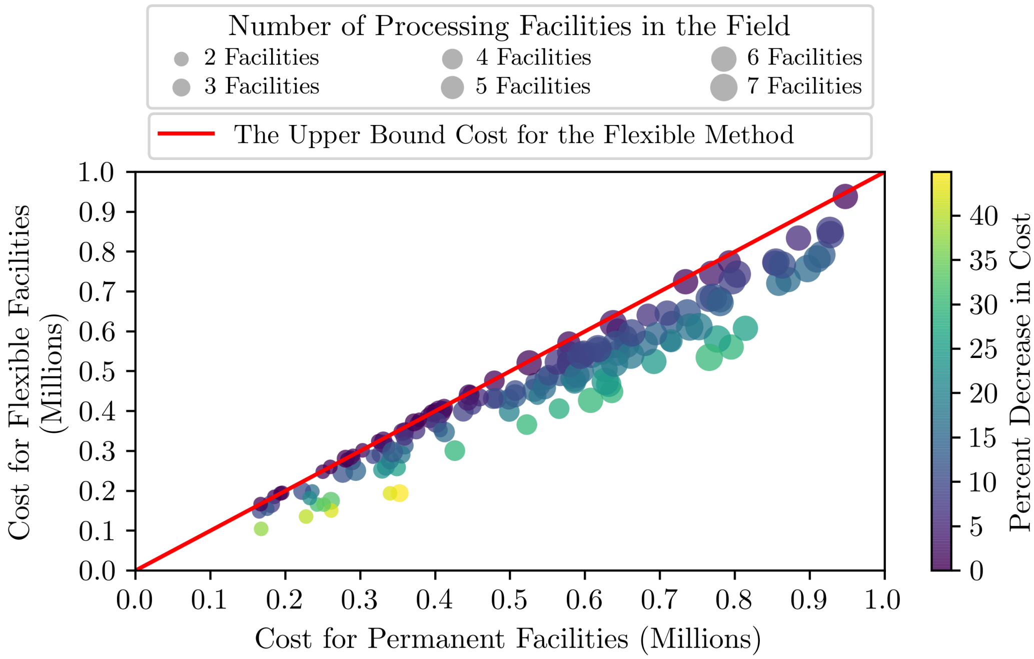


Figure 8.3: Comparison Between Permanent and Flexible Facilities

Given the random production forecasts and the aforementioned parameters we found that our methodology was on average at least 12% more cost effective than utilizing permanent facilities. As the number of facilities in the field increased so did our advantage: (g.i) for two facilities our advantage is 9.5% on average; (g.ii) for three facilities the advantage is 12.0% on average; (g.iii) for four facilities the advantage is 12.0% on average; for five facilities the advantage is 12.2% on average; for six facilities the advantage is 13.6% on average; and for seven facilities the advantage is 13.6% on average

9. CONCLUDING REMARKS

In this work, we have presented a methodology that utilizes multi-stage stochastic programming to aid decision makers in solving multi-facility capacity planning problems for oilfield infrastructure planning. We have postulated, that it is more effective on average to utilize facilities that are composed of transportable processing units that can be reallocated between facilities with flexible capacities as opposed to permanent facilities with fixed capacities. The approach we presented for solving this problem incorporates a novel recourse function, which allows the decision maker to quantify the effect of postponing production to a later time period. We illustrated our methodology through the use of a motivating example, where the decision maker is an exploration and production company trying to develop infrastructure in an unconventional natural gas field. Through the use of this motivating example, we have highlighted how our approach can improve the flexibility the decision maker has in combatting uncertainty. We have shown that in the worst case scenario utilizing flexible facilities is equivalent to utilizing permanent facilities, since each flexible facility is comprised of one skid that is never relocated to a different facility. With that said, we have illustrated, through the use of random test cases for our motivating example, that our methodology is on average 12% more cost effective that utilizing permanent facilities.

REFERENCES

[1] *Annual Energy Outlook 2017 with Projections to 2050*. US Energy Information Administration, US Department of Energy, DOE/EIA, 2017.

[2] M. G. Drouven and I. E. Grossmann, "Multi-period planning, design, and strategic models for long-term, quality-sensitive shale gas development," *AIChE Journal*, vol. 62, no. 7, pp. 2296–2323, 2016.

[3] L. Mattar *et al.*, "Production analysis and forecasting of shale gas reservoirs: case history-based approach," in *SPE Shale Gas Production Conference*, Society of Petroleum Engineers, 2008.

[4] A. Aseeri, P. Gorman, and M. J. Bagajewicz, "Financial risk management in offshore oil infrastructure planning and scheduling," *Industrial & Engineering Chemistry Research*, vol. 43, no. 12, pp. 3063–3072, 2004.

[5] D. C. Cafaro and I. E. Grossmann, "Strategic planning, design, and development of the shale gas supply chain network," *AIChE Journal*, vol. 60, no. 6, pp. 2122–2142, 2014.

[6] B. Wang, Y. Liang, J. Zheng, T. Lei, M. Yuan, and H. Zhang, "A methodology to restructure a pipeline system for an oilfield in the mid to late stages of development," *Computers & Chemical Engineering*, vol. 115, pp. 133–140, 2018.

[7] H. Luss, "Operations research and capacity expansion problems: A survey," *Operations Research*, vol. 30, no. 5, pp. 907–947, 1982.

[8] J. C. Bean, J. L. Higle, and R. L. Smith, "Capacity expansion under stochastic demands," *Operations Research*, vol. 40, no. 3-supplement-2, pp. S210–S216, 1992.

[9] P. K. Naraharisetti and I. A. Karimi, "Supply chain redesign and new process introduction in multipurpose plants," *Chemical Engineering Science*, vol. 65, no. 8, pp. 2596–2607, 2010.

[10] S. D. Jena, J. Cordeau, and B. Gendron, "Dynamic facility location with generalized modular capacities," *Transportation Science*, vol. 49, no. 3, pp. 484–499, 2015.

[11] M. T. Melo, S. Nickel, and F. S. Da Gama, "Dynamic multi-commodity capacitated facility location: a mathematical modeling framework for strategic supply chain planning," *Computers & Operations Research*, vol. 33, no. 1, pp. 181–208, 2006.

[12] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming.* Springer Science & Business Media, 2011.

[13] I. E. Grossmann, R. M. Apap, B. A. Calfa, P. García-Herreros, and Q. Zhang, "Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty," *Computers & Chemical Engineering*, vol. 91, pp. 3–14, 2016.

[14] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on Stochastic Programming: Modeling and Theory.* Society for Industrial and Applied Mathematics, 2009.

[15] B. Defourny, D. Ernst, and L. Wehenkel, "Multistage stochastic programming: A scenario tree based approach to planning under uncertainty," in *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions*, pp. 97–143, IGI Global, 2012.

[16] M. M. El-Halwagi, *Sustainable Design Through Process Integration: Fundamentals and Applications to Industrial Pollution Prevention, Resource Conservation, and Profitability Enhancement.* Elsevier, 2011.

[17] *Gurobi Optimizer Reference Manual, Version 8.0.* Gurobi Optimization, 2018.

APPENDIX A

DATA SET

Table A.1: Data for the Implicit Scenario Graph for the Case Study

Node	Time Period	f_a	f_b	f_c	f_d	Probability	Forecast
n_1	1	100	0	0	0	1	1, 2, 3, & 4
n_2	2	95	45	0	0	0.05	1 & 2
n_3	2	85	45	0	0	0.95	3 & 4
n_4	3	90	45	30	0	0.05	1
n_5	3	90	40	30	0	0.95	2
n_6	3	75	45	30	0	0.05	3
n_7	3	75	40	30	0	0.95	4
n_8	4	85	50	80	0	1	1
n_9	4	85	35	80	0	1	2
n_{10}	4	60	50	80	0	1	3
n_{11}	4	60	35	80	0	1	4
n_{12}	5	95	60	75	20	1	1
n_{13}	5	95	30	75	20	1	2
n_{14}	5	50	60	75	20	1	3
n_{15}	5	50	30	75	20	1	4
n_{16}	6	130	55	65	50	1	1
n_{17}	6	130	20	65	50	1	2
n_{18}	6	45	55	65	50	1	3
n_{19}	6	45	20	65	50	1	4

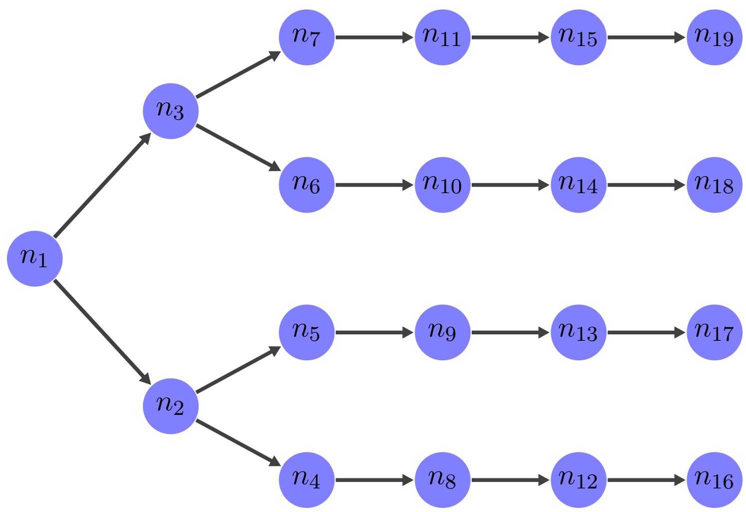


Figure A.1: Implicit Scenario Graph for the Case Study